Commonwealth of Massachusetts Committee on Information Technology

Logical Network Architecture Guidelines

Version 2.0
July 1998

Presented by the Strategic Planning Bureau of the Information Technology Division

Table of Contents:

- Executive Summary
- Year 2000 Compliance
- Logical Architecture

- Overview
- Component Network Architecture
- Division of Management Responsibilities
- Connecting to MAGNet
- Protocol Architecture

- Link Layer (OSI Layer 2) Connections

- LAN Emulation (LANE)

- Network Layer (OSI Layer 3) Routing Protocols and the MAGNet Backbone
    - OSI
    - TCP/IP
    - LAN Protocols
    - Vines IP
    - IPX
- Addressing
    - Addressing Responsibilities
    - IP Address Plan
    - IPV6 Guidelines
    - IP Traffic Filtering
    - Implementing the Commonwealth IP Address Plan

Executive Summary

This document is a guideline for the network architecture when establishing or updating agency networks within the Commonwealth of Massachusetts and its agencies. The foundation of this architecture is the MAGNet (Massachusetts Access to Government NETwork) wide area network and its constituent multiprotocol routers and ATM switches. This is a multiprotocol network, supporting a wide range of heterogeneous systems and protocols used in the Commonwealth's IT environments.

MAGNet has the advantage of offering cost-effective wide-area connectivity utilizing leveraged telecom contracts and providing the opportunity for inter-agency data exchange and access to shared resources.

This standard calls for convergence on a single protocol standard (TCP/IP). A single, open, robust networking protocol charts a path toward greatly increased connectivity and widespread interoperability between the many systems used by the Commonwealth and its agencies.

The purpose of this document is to set standards and provide advice for state agencies when making decisions about how to construct or evolve their networks to be capable of leveraging the MAGNet telecommunications infrastructure. This document is not intended to solve specific implementation problems but to act as a set of guidelines that may be used to maximize interoperability between the networks and computer systems of various agencies and save on telecommunication costs.

Year 2000 Compliance

As directed by the Commonwealth of Massachusetts Executive Office for Administration & Finance on September 29, 1997, all purchases by Commonwealth agencies of new software, systems, enhancements or equipment shall be Year 2000 compliant. Information on the Commonwealth's Year 2000 Project is available on the Web at http://www.state.ma.us/y2k/.

The Information Technology Division, through its Year 2000 Project Management Office, will continue to offer assistance to agencies in their Year 2000 compliance efforts. The Operational Services Division is available to assist with technology procurement matters related to Year 2000 compliance.

Logical Architecture

Overview

The foundation of the Commonwealth's logical architecture is the multiprotocol routers and ATM switches. Although the Commonwealth's enterprise network has been designed to accommodate multiple protocols, it is the goal of the architecture to converge the enterprise on one open standard. **The major target protocol standards for this architecture are TCP/IP and the OSPF routing protocols.** Convergence on these standards is an important step toward interoperability between systems across every organization in the Commonwealth. However, the migration to these standards cannot be done overnight. Thus, the Commonwealth has been executing the following two-step strategy toward interoperability since 1994 to create a multiprotocol backbone:

- **Step 1**: get all end systems connected to a common enterprise network that can accommodate protocols that are widely used today and embedded in systems which the Commonwealth and its agencies cannot afford to re-engineer today.
- **Step 2**: decrease the number of protocols supported on the backbone as end systems are converted to an open systems standard TCP/IP.

A multiprotocol enterprise backbone is a transition/migration tool toward an open single protocol backbone. At first, the existing protocols must be supported by the backbone, as more and more end systems connect. Then proprietary protocols such as SNA, and Vines IP, etc. will disappear as systems are replaced with more open ones. With everyone connected to a common enterprise network, it will be possible to control the proliferation of protocols and gradually enforce restrictions on the use of non-standard protocols.

Reducing the number of protocols carried by an enterprise backbone serves two important purposes. First, reducing the number of protocols decreases the complexity of the enterprise network, making it easier and less costly to manage and making it more reliable. Secondly, it makes more inter-system communication possible, thereby allowing sharing of data. Additionally, reducing the number of protocols on the enterprise backbone can improve the performance of the backbone routers.

Network architectures can be described from several perspectives. This document describes the network architecture from a physical component view and a protocol view.

Component Network Architecture

The component network architecture for the Commonwealth of Massachusetts is depicted in Figure 1. It consists of a core backbone network, called MAGNet, and attached agency networks. MAGNet is composed of an ATM Metropolitan Area Network (MAN) interconnected with a wide area network (WAN) made up of T1 and frame relay links. Over a period of time, more of the T1 and frame relay links are being replaced by ATM connections as the ATM MAN is expanded to cover more of the state.

(NOTE: Figure 1 illustrates the components and the interfaces found in the Commonwealth's network. It does not illustrate the actual design of the network. Such an illustration, showing the redundant components and links and complex relationships are beyond the scope of this document.)

Agencies interface to MAGNET at the "Service Access Point" (SAP), as indicated in Figure 1. The SAPs indicate the line of demarcation between components of the network managed by ITD CSB and the components managed by individual agencies. For convenience, components on the MAGNET side of the SAP are identified as "Core Hubs" (CH), and those components on the agency side of the SAP are identified as "Group Hubs" (GH).

Division of Management Responsibilities

As indicated in Figure 1, agencies interface to MAGNET at the "Service Access Point" (SAP). The SAPs indicate the line of demarcation between components of the network managed by ITD CSB and the components managed by individual agencies. For convenience, components on the MAGNET side of the SAP are identified as "Core Hubs" (CH), and those components on the agency side of the SAP are identified as "Group Hubs" (GH).

The management of the logical network is shared between individual state agencies and the Information Technology Division's Communications Service Bureau (ITD CSB). Agencies have the responsibility of managing the logical network from the end-systems to and including the GH, while ITD CSB manages the logical network from and including the CH into the wide-area. The links between the GH and the CH may fall into a shared management domain, but is usually the responsibility of the agency. Large executive agencies with internetworking expertise may negotiate with ITD CSB to manage this link, or at their discretion, request that this link be managed for them. It is, however, in the best interest of individual agencies to have ITD CSB manage the GH-CH links wherever possible. The GH-CH connections are stable and will not change as frequently as the GH-ES links. Agencies operating with limited resources probably do not have the personnel or the training dollars to take on this responsibility.

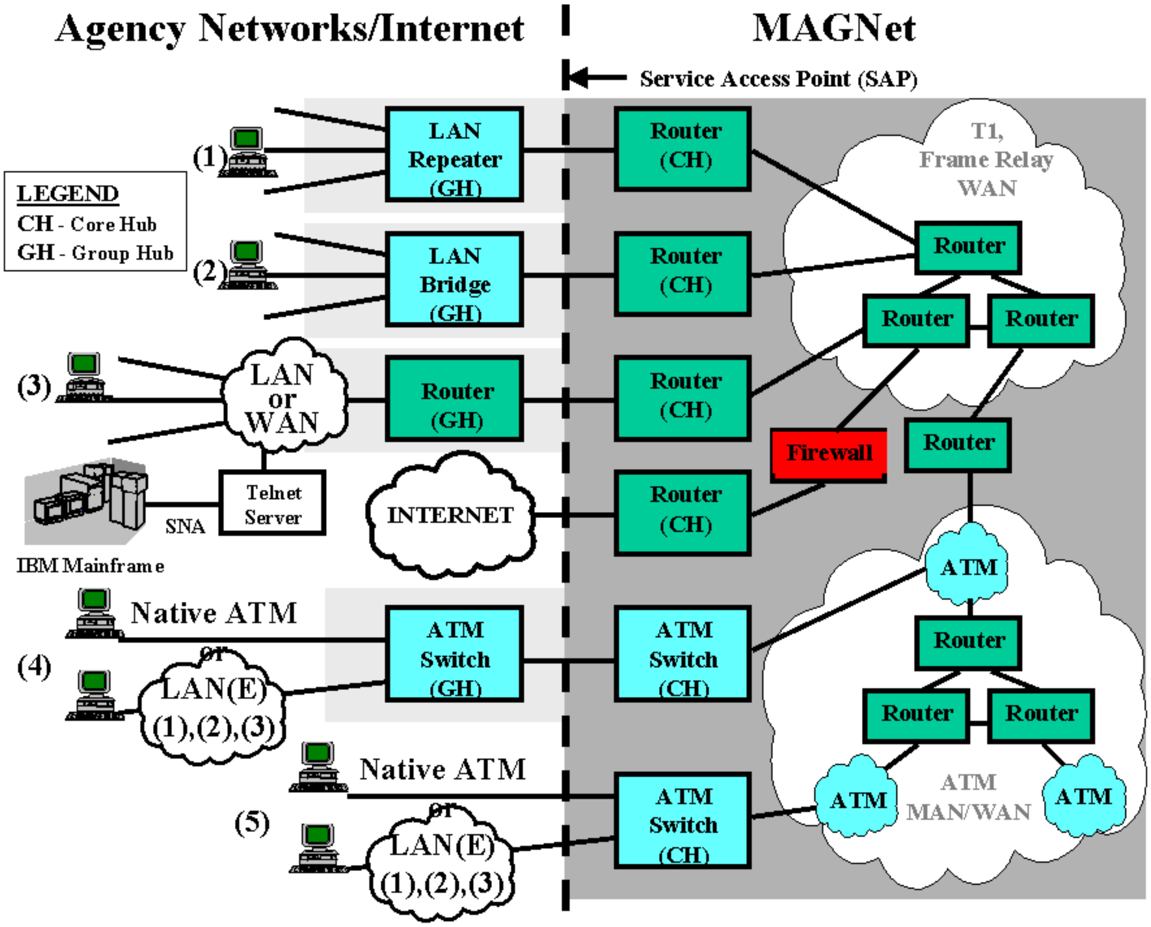

Figure 1: MAGNet Component Architecture

Figure 1 illustrates the separation in logical network responsibilities. ITD CSB normally manages components to the right of the SAP, and those to the left of the SAP are normally managed by the agencies.

ITD CSB establishes the guidelines for the protocols that may traverse between CHs in the MAGNet backbones while individual agencies have the freedom to choose the logical networks they will support and use. For those protocols, which will be used on an enterprise scale, ITD CSB has the responsibility for naming and address policy/distribution. ITD CSB has acquired two Class-B IP address for the Commonwealth and all of its agencies. ITD CSB has the responsibility for providing blocks of addresses to individual state agencies, which in turn have the responsibility for assigning addresses to particular network devices.

Connecting to MAGNet

The methods that the agencies use to connect to MAGNet are illustrated in Figure 1. The method of connection differs depending on whether they are connecting to the ATM portion of the backbone, or to the T1/Frame Relay portion of the network. Agencies located within buildings serviced by the ATM network may connect to the ATM network by means of Ethernet ports on an ATM building group hub as illustrated in Figure 1. Agencies not serviced by the ATM network must connect via IP routers attached to the MAGNet T1/Frame Relay links.

The principal determinant of which interconnect technology agencies should be used to link the GH and CH is the LAN protocol environment. Figure 1 illustrates the major alternatives for interconnecting the GH and CH.

Scenario 1: The CH is a multiprotocol router and the GH the agency uses to connect to MAGNET is a simple LAN repeater. This is appropriate when an agency connects a few moderately sized workgroups using a single network layer protocol. Small sites might have a single LAN segment connected to the MAGNet Core Hub.

Scenario 2: The CH is a multiprotocol router and the GH the agency uses to connect to MAGNET is a LAN bridge. This is appropriate when an agency uses multiple network layer protocols. The use of filtering bridging will provide some traffic isolation, effectively increasing the bandwidth available for each LAN on a GH. The advantage of using bridging in this situation is that network layer protocols will be transparent to it. The physical architecture specifies Ethernet as the standard LAN media. Transparent Spanning Tree bridging is therefore the bridging standard for the architecture. Medium size sites might require several LAN segments.

The Commonwealth has little investment in Token Ring technology, and this is advantageous. Token Ring is more expensive than Ethernet, and SRB (Source Route Bridging) technology is not as efficient as the Spanning Tree and Transparent bridging used in Ethernet LANs. Where combinations of Token Ring and Ethernet LANs exist together within state agencies, SRT bridging will be used to connect the Ethernet and Token Ring worlds. Network managers are discouraged from the purchase of new Token Ring and SRT technology.

Scenario 3: The CH is a multiprotocol router and the GH the agency uses to connect to MAGNET is a multiprotocol router. This is appropriate when the agency implements multiple routable protocols and traffic isolation between GH workgroups is required in a LAN environment, or the agency network is a multiprotocol WAN. Large sites may require several LAN segments. Routing is highly recommended to connect them to the MAGNet Core Hub. This affords traffic isolation, fault containment, limited broadcast and multiple paths.

If the agency uses non-routable protocols within a LAN environment, then bridging is also required. The GH may be a "brouter" (an integrated bridge/router). The GH brouter will route all routable protocols, and bridge non-routable protocols. Brouters are complex devices since they enable internetworking by making it possible to interconnect almost every kind of end system. The resulting complexity of the brouted network environment makes it very difficult to optimize the network configurations. Whenever an agency believes it may need to investigate brouter technology, it should seriously evaluate whether the protocol environment may be simplified through migration to all routable protocols, or whether the use of routable and non-routable protocols may be isolated to physically separate GHs, enabling both a GH bridge and a GH router to handle the appropriate protocols separately.

Note that this scenario also illustrates the interconnection of the mainframes using SNA to magnet via a TELNET server. Agency workstations running the TN3270 protocol emulation software and a TCP/IP protocol stack can communicate with the mainframe over MAGNet. Agencies must make every effort to connect to MAGNet utilizing TN3270.

Scenario 4: The CH is an ATM switch and the GH the agency uses to connect to MAGNET is also an ATM switch. This is appropriate if the agency already has an ATM network, or is using a carrier provided ATM service, and wishes to connect to MAGNet. MAGNet will provide the necessary routing to get to other points in the network. It may do this using three methods:

- By creating ATM layer 2 direct links to destinations that participate in the ATM portion of the network
- By using network layer 3 to route the data to its destinations
- By using a combination of ATM layer 2 routing (to get to a network layer 3 router) and network layer 3 routing to get to the destination.

Note that the agency may connect to the GH ATM switch using either native ATM or standard LAN protocols (Ethernet, IEEE 802.3,5). End-systems using standard LAN protocols must use LAN emulation (LANE) to communicate over the ATM network.

Scenario 5: The CH is an ATM switch and the agency uses either native ATM or standard LAN protocols (Ethernet, IEEE 802.3,5) to connect to MAGNet. This is the method used by most agencies to connect to the ATM portion of MAGNet. MAGNet will provide the necessary routing to get to other points in the network. It differs from Scenario 4 only in that the agency does not already have an ATM network.

Internet Connectivity: Magnet connects to the Internet via an IP router backed up by a firewall. This connection provides access for all Commonwealth agencies to the Internet.

While it is preferable, over the long term, to use MAGNet's routing capabilities to control and integrate traffic among and between every state agency, the use of MAGNet as only a shared, multiplexed transmission infrastructure is acceptable for legacy applications in the near term. The Commonwealth's investment in SNA makes it unreasonable not to continue to support SNA users by providing multiplexed access to MAGNet. It is unlikely that the Commonwealth or its agencies will secure funding to completely integrate the large SNA leased line installed base into a router-based internetwork in the near future. Agencies should make every effort to migrate to a TN3270 TCP/IP based access method.

Protocol Architecture

While the above section provided a high level overview of the physical and component architecture for the Commonwealth's enterprise network, this section will provide a detailed explanation of the logical network which rides on top of the physical framework. The logical network exists in communications software and consists primarily of layers 2, 3 and 4 in the OSI model. In this section we will discuss protocols, addressing and routing. Logical networks will and can change freely without disrupting/modifying or altering the physical network configuration.

The logical network architecture identifies the logical connections over which traffic may flow, the routing technology used to determine the specific connections over which data may flow, and the data communication protocols used for the data transfers. The logical architecture is constrained by the physical component architecture, e.g., logical connections can only occur over physical links.

Link Layer (OSI Layer 2) Connections

The Commonwealth enterprise network is based on connectionless internetwork layers (e.g., IP and IPX). Data flows within the network layer must flow over logical connections established at the link layer; e.g., Frame Relay or ATM virtual circuits (VCs). Connection establishment at the link layer may be performed either through management actions, e.g., permanent virtual circuits (PVCs), or dynamically upon end-user initiation, e.g., switched virtual circuits (SVCs).

The ATM portion of MAGNet contains a logical configuration of permanent virtual circuits that determine the ATM link topology. Where ATM is present in buildings that house the agencies, the logical topology formed by the VCs can be a full mesh that connects every ATM port to another ATM port to which an agency LAN or end-system connects. It can also be a full mesh that interconnects all ATM ports within an agency. Such a mesh enables communications between all end-stations attached to the mesh without network layer routing. The Commonwealth also uses LAN Emulation (LANE) with ATM, which provides each mesh with the characteristics of a broadcast LAN. Such a LANE enabled mesh of ATM VCs is called a virtual LAN (VLAN). Communication between VLANs can occur at layer 3 using network layer routers.

The Commonwealth has constructed ATM VLANS for each agency so as to eliminate network layer routing within an agency, and provide traffic isolation between agencies. Furthermore, agency VLANs are interconnected using network layer multiprotocol routers.

LAN Emulation (LANE)

Most data traffic in agencies is sent over Local Area Networks (LANs), such as Ethernet/IEEE 802.3 and IEEE 802.5 networks. The services provided by today's LANs differ from those of ATM, for example:

1. The messages may be characterized as connectionless, versus the connection-oriented approach of ATM;

2. Broadcast and multicast are easily accomplished through the shared medium of a LAN;

3. LAN MAC addresses, based on manufacturing serial numbers, are independent of the network topology.

In order to use the vast base of existing LAN application software, the ATM Forum has defined an ATM service, herein called "LAN Emulation," that emulates services of existing LANs across an ATM network and can be supported via a software layer in end systems.

The LAN Emulation service enables end systems (e.g. workstations, servers, bridges, etc.) to connect to the ATM network while the software applications interact as if they are attached to a traditional LAN. Also, this service supports interconnection of ATM networks with traditional LANs by means of IEEE bridging methods. This allows interoperability between software applications residing on ATM-attached end systems and on traditional LAN end systems.

Customers expect to continue to use existing LAN applications with ATM. The LAN Emulation service has proven to be important to the acceptance of ATM, since it provides a simple and easy means for running existing LAN applications in the ATM environment.

LAN Emulation defines a MAC service emulation, including encapsulation of MAC frames (user data frames). This approach to LAN emulation provides support for the maximum number of existing applications.

The following LAN-Specific Characteristics are provided by LANE:

- Connectionless Services

  LAN stations today are able to send data without previously establishing connections. LAN Emulation provides the appearance of such a connectionless service to the participating end systems.

- Multicast Services

  The LAN emulation service supports the use of multicast MAC addresses (e.g. broadcast, group, or functional MAC addresses). The need for a multicast service for LAN Emulation comes from classical LANs where end stations share the same media.

- MAC Driver Interfaces In ATM Stations

> The main objective of the LAN emulation service is to enable existing applications to access an ATM network via protocol stacks like APPN as if they were running over traditional LANs. Since in today's implementations these protocol stacks are communicating with a MAC driver, the LAN emulation service has to offer the same MAC driver service primitives, thus keeping the upper protocol layers unchanged.

Network Layer (OSI Layer 3) Routing

Implicit in the design of the enterprise network architecture is a conversion to a routed architecture for the core of the Commonwealth's enterprise internet. MAGNet is the result of the Commonwealths conversion to a routed architecture. Experience with a bridged core approach to LAN interconnect throughout the industry has shown that it is unreliable and not robust. An architecture based on routing, although not perfect, will be an improvement. Routing is the key step toward stable logical network operation mainly due to the following:

- Subnetworks are created, providing boundaries between LANs (and VLANs) and WANs which restrict the propagation of broadcast storms which can cripple a large internet.
- A wider range of physical topologies and WAN services are accommodated.
- A range of congestion management options are available and will continue to be developed.
- New routing algorithms, which converge faster after a change in topology occurs, are being developed for an ever increasing range of protocols.

At the heart of the logical network is routing technology and data communication protocols such as IPX, TCP/IP and OSI. The logical architecture is a multiprotocol enterprise backbone implemented by multiprotocol routers as typified by MAGNet's Core Hubs. The typical Group Hub, a workgroup-class hub, will not normally provide OSI Layer 3 routing. The majority of existing GHs are LAN repeaters (Ethernet hubs) or bridges connected to the CH. **The use of software-based routing in servers is not recommended**.

Routing is the optimal technology for interconnecting GHs and CHs, particularly in large and/or multiprotocol LAN environments. Although only a few agencies currently have complex workgroup LANs which warrant a router at the GH level, routing's advantages make it the most desirable technology for the long term. Router vendors have simplified routing, making it easier to deploy and manage, even at the GH level. However, agencies will continue to use bridging within small-to-medium size LANs for some time to come. Inexpensive, transparent filtering bridges will be easier to install and manage for most state agencies.

The ability to quickly adapt to changes in topology caused by link failures and load changes due to changing traffic flows is critical to the functioning of the connectionless network layer. Static routing algorithms will not satisfy this adaptive requirement.

Instead, routing is performed on the basis of metrics, e.g., link load, and the number of hops to the destination. These metrics must be updated as the network topology or traffic load changes. Routing protocols are used between routers in the network to provide the data necessary to dynamically update the metrics upon which routing is based. The dominant dynamic routing protocol was RIP. It has some shortcomings that OSPF (Open Shortest Path First) later alleviated. **OSPF is now the IP dynamic routing protocol of MAGNet**.

When routers are used at the GH-CH interface for the TCP/IP protocol stack, they must use the OSPF routing algorithm. Routable, proprietary protocols such as IPX require their own routing algorithms. IPX uses IPX RIP, RIP II or Enhanced RIP dynamic routing protocols. Network Operating System Vendors (NOS) vendors are converging on the TCP/IP protocol stack as a transport standard fully supported alongside their own native protocol stacks. Over time, NOS vendor's support of TCP/IP will make it possible to use OSPF only as the one standard routing algorithm at even the workgroup level of the enterprise.

Protocols and the MAGNet Backbone

TCP/IP will only be supported in the MAGNet backbone without providing a migration strategy. The network protocols currently supported on MAGNet are IP, IPX/SPX, and Vines IP.

The network protocols currently supported by agencies throughout the commonwealth are a mixture of open and proprietary protocols.

Open Protocols

OSI

The Open Systems Interconnection (OSI) communication protocols are not supported within the Commonwealth's Enterprise Network Architecture. OSI never achieved the widespread deployment originally anticipated due to the onslaught of the Internet.

TCP/IP

Only TCP is fully supported within the Commonwealth's Enterprise Network Architecture. TCP is the enterprise standard, and is considered open by virtue of its continued development overseen by the IETF. Depending on the rate of change in the application environment, as time passes a heavier weight will be assigned to the TCP protocol as the backbone protocols supported by the ITD CSB. The enterprise network will converge on this protocol toward the end of this planning horizon.

The intention of these standards is to encourage the widespread use of the TCP/IP protocol, especially for server-to-server communications. Use of TCP/IP to the desktop, while not discouraged, may not be cost-effective for some agencies at this time and might be provided through TCP/IP services provided by the LAN server.

Proprietary Protocols

TCP/IP Gateway for SNA: The TN3270 standard is the best alternative available for many enterprises with a significant investment in both SNA SDLC-based mainframe applications and PC LANs interconnected via TCP/IP backbones. The Commonwealth should migrate as many systems as is cost justifiable to the TN3270 standard.

The TN3270 gateway standard provides a mechanism for delivering terminal/host access to existing SNA application via a TELNET server and the TCP/IP protocols across an internet backbone. The advantage of using TN3270 is that existing mainframe applications need not be changed, thereby preserving the Commonwealth's application investment while bringing the use of such applications into compliance with this logical architecture. The majority of IBM mainframe systems may be run using TN3270. It supports terminal types 2, 3, 4, and 5. Several vendor implementations support multiple concurrent client sessions, file transfer and also provide printer emulation/support to DOS PC environments.

The TN3270 approach allows FEPs, terminal cluster controllers, and the SNA protocol to be replaced with TELNET servers, and a router-based internetwork using the TCP/IP protocol. The use of SNA in the network is limited to the connection between the host and the TELNET server to which it is attached. The TELNET server, in turn, connects via the IP-based router backbone, to PC clients running the TN3270 emulation software and a TCP/IP protocol stack.

LAN Protocols

The Commonwealth Standards Subcommittee has issued a document entitled Standards and Guidelines for Local Area Networks (LANs) which strongly recommends that individual state agencies deploy NT for new installations. The document currently requires that all departments with servers at either Banyan Vines version 4.x - 7.x must upgrade to Vines version 8.5 no later than December 31, 1998. Novel Netware Version 4.1 with TCP/IP Option(s) such as Novix or LAN Workplace may also be retained. These two LANs natively operate on the Vines IP and IPX protocols. However, servers must also support routing of native TCP/IP traffic to and from administrative desktops. Agencies should reference the latest version of Standards and Guidelines for Local Area Networks for updates. The use of each in the MAGNet environment is discussed below.

Vines IP

Banyan provides both server-server and client-server mechanisms for integrating PC LAN workgroups into a TCP/IP network:

The TCP/IP server-server option will be the most useful in the Commonwealth's enterprise network. We recommend that TCP/IP be used as the Vines server-server communication standard wherever possible, but specifically for server-server communications via MAGNet. Depending on the requirements of individual sites, only a single Vines server will be required to act as the TCP/IP gateway to MAGNet, meaning that all clients of other servers at that site may use the TCP/IP gateway services of a single Vines gateway server. However, a single gateway server may become a bottleneck for sites which increasingly require connections over the wide area.

IPX

Novell's IPX was designed primarily for LAN applications and has deficiencies in its support for wide area networking. These include little congestion management, no load balancing and superfluous messages which flood the wide area. IPX has no sliding window mechanism to control the flow of messages between client and server. The NCP, or Netware Core Protocol(s), allow only a single outstanding message to exist between client and server at any time. Each message must receive an acknowledgment before the next message may be sent. When multiple bridge and router hops or WAN circuits and switches are involved, the delay between packets becomes a serious problem. Additionally, IPX uses a protocol called Service Advertisement Protocol (SAP) which broadcasts to every server on the network the availability of services to Netware clients. This broadcast occurs every 30 seconds and becomes more troublesome when the network grows in size and extends over the wide area.

Addressing

With a routed core network architecture as the enterprise backbone, an addressing plan for each protocol is an absolute requirement. Since routing will link the Commonwealth agency's LANs into an enterprise network, a large number of subnetworks will be created, increasing the need for a coherent addressing plan. Further, since the backbone is multiprotocol, every routed protocol on the network requires an addressing plan. The two principal internetworking protocols used by the Commonwealth end systems are IPX and IP. Other network protocols do exist, but are being deprecated in favor of these. IP will become increasingly prevalent such that ultimately all network entities will have IP addresses if only for the purposes of remote administration and operation.

Addressing Responsibilities

ITD CSB has registration authority for two blocks of class B IP addresses. As such, it has the overall responsibility to develop an addressing policy and distribute IP address space blocks to the agencies for their use. Also, the industry standard practice for assigning IPX addresses in an IP-dominated network is to represent them as the hexadecimal equivalent of the IP address allocated to an interface of each network entity. **It is recommended that the Commonwealth follow this practice for IPX address assignments. Thus, when an IP address block would be allocated to an agency, the equivalent IPX addresses would be included in that allocation**

ITD CSB's addressing responsibilities include:

- Develop a strategy for partitioning blocks of addresses to individual agencies for each protocol.

- Distribute blocks of addresses to each agency for each protocol.
- Develop an addressing domain of subnetworks within the MAGNet backbone for each protocol.
- Maintain an addressing database which captures all address blocks belonging to agencies for each protocol in use.
- Develop a cut over strategy for changing subnetwork address domains. A scenario for this activity is when a subnetwork outgrows its addresses space and must be either split in two subnetworks or traded with another less populated subnetwork.

For TCP/IP and OSI addresses, it is imperative that addresses be registered with official registration authorities to insure their global uniqueness. This will be essential in light of the increasing use of inter-enterprise internetworking, and the continuing need for enhanced interconnection and cooperation between state, local, and federal government with the private sector.

IP Address Plan

The classification and structure of an IP address, with its {net, subnet, host} hierarchy, are deceptively simple. However, within the IP routing paradigm addresses are more than simply locators — they can embody wider aspects of an enterprise’s character, from geographic dispersion to operating procedures, administrative responsibilities, security policy, and projected growth. Because addresses take on these implicit meanings, the effort to change addresses or even subnet masks in an operational environment is not trivial, and without due diligence can be severely detrimental to the business activities supported by the network. Thus, it is vitally important that the Commonwealth establish and adhere to a formal address plan, both to ensure consistency in addressing and to prevent addressing failures as the network evolves.

IP Subnet and Host Addressing Considerations

In order to determine a subnet mask, the network administrator must estimate how many hosts each subnet is expected to have. As it is often impossible to predict how large any given subnet will become, subnet sizes are often overestimated to obviate the need for subsequent renumbering. This inefficiency is further compounded by routing protocols that do not support variable length subnet masks (VLSM), and thus require every subnet of a given network number to be allocated a host address space equivalent to that of its largest subnet.

To address the problem of size overestimation, RFC-1219 defines a "mirror image" counting method for subnet numbering. In essence, this strategy allows subnet numbers to expand from most significant bit towards least significant bit, and host numbers to expand from least significant bit to most significant bit in the "shared" space of the (unmasked) host field as illustrated in Figure 2. This technique contrasts favorably to the more commonly-practiced approach of pre-determining the subnet mask and numbering the subnets from least to most significant bit within that mask—such an approach entails significant renumbering whenever any one of the subnets exhausts its assigned host address space, even if the assigned number of subnets is (often significantly) fewer than permitted by the subnet mask. By complying with RFC-1219, the available number of subnets and hosts per subnet may be adjusted by changing the subnet mask, without having to amend the addresses that have already been assigned.

However, it should be recognized that changing subnet masks in an operational network is a non-trivial administrative procedure (albeit significantly less disruptive than renumbering hosts). Moreover, the full benefits of RFC-1219 are predicated on the use of VLSM.

Subnet sizing also has implications for network infrastructure services that the Commonwealth must consider in the detailed design of its address plan:

- DNS Zone Delegation. Two general precepts of a Domain Name System (DNS) design is that names/addresses be translatable in either direction, and that internal zones may be created and delegated to independent administrative authorities. Unfortunately, PTR records in the in-addr.arpa domain that are used for reverse (address-to-name) lookups interpret IP addresses much as text strings that are delimited on conventional "dotted decimal" (octet) boundaries. Consequently, delegating name authority independently from address block assignment, and the use of subnetting along non-octet-aligned boundaries, can have significant impact on the complexity (and hence support cost) associated with the administration of delegated DNS zone databases

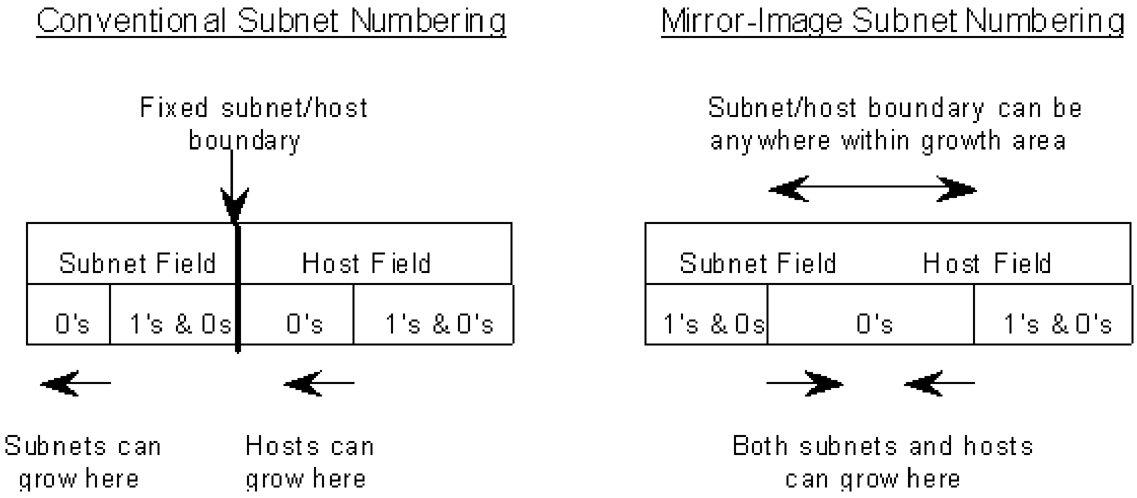
Conventional Subnet Numbering
Fixed subnet/host boundary
Subnet Field
Host Field
0's
1's & 0s
0's
1's & 0's
Subnets can grow here
Hosts can grow here
Mirror-Image Subnet Numbering
Subnet/host boundary can be anywhere within growth area
Subnet Field
Host Field
1's & 0s
0's
1's & 0's
Both subnets and hosts can grow here

Figure 2. Subnet Number Assignment

IP Address Allocation

ITD CSB has registration authority for two blocks of Class B IP addresses assigned by IANA. However, since the secure perimeter surrounding the Commonwealth network fundamentally isolates all internal addresses from external access, the Commonwealth address plan could, in principle, be based on any convenient set of network numbers. Should the Commonwealth run out of its allocated Class B IP address space, it could assign addresses in accordance with RFC 1918. RFC-1918, "Address Allocation for Private Internets", defines the following address space that is reserved for networks that will never communicate directly with the global Internet:

- 1 Class A network number (10/8 prefix)
- 16 contiguous Class B network numbers (172.16/12 prefix)
- 256 contiguous Class C network numbers (192.168/16 prefix)

The reserved network numbers listed above represent a more-than-adequate address space for the Commonwealth, and, as the IAB's recommended "best practice", are well known and therefore explicitly filterable if inadvertently leaked. Consequently, while the selection of specific network numbers and subnet masks is a design task beyond the scope of this architecture, the Commonwealth network address plan could be based on RFC-1918 for its internal networks. However, external "DMZ" networks are publicly accessible and must therefore use IANA-registered network numbers.

IPV6 Guidelines

The adoption of IP Version 6 (IPV6) by the Commonwealth is not recommended at this time. IPV6, currently an Internet Proposed Standard, has been proposed to replace the current IP Version 4 (IPV4). However, most proposed enhancements incorporated into IPV6 also work with IPV4, e.g., IP Security. The major enhancement over IPV4, the extended address space, is not needed by the Commonwealth at this time but may be needed in the future. Standard methods have been devised to extend the lifetime of the IPV4 address space, including the use of the IPV4 private address space, with or without address translation, as described in RFC1918, "Address Allocation for Private Internets", and RFC1631 "The IP Network Address Translator (NAT)". Finally, IPV6 seems to be making little headway into vendor products, and it's adoption within the Internet community as the dominant standard is not assured.

IP Traffic Filtering

IP addresses can be structured to help manage network traffic flows when, for example, groups of networks, or of hosts sharing a network, represent closed communities that should not be permitted to intercommunicate. Such communities can readily be identified by reserving specific bits in the subnet or host fields of the addresses, respectively. Routers can then be configured to mask source and/or destination packet addresses against these community identifiers and discard unauthorized inter-community traffic.

This scheme may also be adapted to other purposes appropriate to the Commonwealth business units' requirements; e.g., to designate classes of hosts whose traffic the network routers should service with a given priority or with "reserved" routes (i.e., policy routing), or to easily distinguish between primary and backup networks. The detailed implementation of such structured addresses will depend on the number, size, and interoperability requirements of the Commonwealth's diverse internal communities, the specifics of which are beyond the scope of this architecture.

Implementing the Commonwealth IP Address Plan

Highly-structured addresses with embedded semantics rapidly become non-intuitive, particularly in conventional dotted-decimal notation. The more bits are loaded with semantic significance the harder it is to assign new addresses correctly, and for an operator to interpret diagnostic messages, route tables, etc. without mechanical assistance. Even the mirror-image counting methodology for subnet numbering described in RFC-1219 should be codified to ensure correct implementation. Spreadsheet-based tools are particularly useful for calculating and interpreting both bit-and-mask and range-based fields, and can readily be customized according to the specifics of the Commonwealth's address plan.

Another key to ensuring the proper implementation of the address plan will be to maintain rigorously a register of all assigned addresses. An up-to-date register is vital to eliminating the risk of address duplication, and is also valuable for asset/inventory management and day-to-day operations. However, it is emphasized that responsiveness should not be compromised for rigor — end users are, in general, empowered to create their own host configuration files and may be tempted to pick ad-hoc values if procedures for obtaining approved addresses are not fast and simple.

Address administration can be simplified if hosts learn their addresses from a server rather than through error-prone manual configuration. Server-assigned addresses are often used in desktop hosts (PCs, Macs, etc.), and are especially common where users have remote access to a network on a dial-in basis. Such addresses are generally temporary and may be automatically reassigned after the host disconnects. However, some applications do not respond well to frequently-changing addresses, especially those registered with the Domain Name System. Products for "Dynamic DNS" and automatically synchronizing DHCP and DNS server databases are becoming available.